RECEIVED P8193
CENTRAL FAX CENTER
MAY 09 2008

REMARKS

Claims 32-35 and 37-56 are pending in the application. Claims have been rejected under 35 U.S.C. §103(a) as being deemed unpatentable over Beshai et al. (U.S. Patent No. 6,721,271), Massa (U.S. Patent No. 6,721,271, Edholm (U.S. Patent No. 6,721,271), Mauger et al. (U.S. Patent No. 6,917,586), Wang (U.S. Patent No. 6,477,612.) and Colgate (U.S. Patent No. 5,594,889). Of the Claims, Claims 32, 40, 44, and 48 are independent. The application as argued herein, is believed to overcome the rejections.

Regarding Rejection of Claim 36 under 35 U.S.C. §103(a)

The Office indicates in the Office Action Summary that Claims 32-35 and 36-56 are pending. In the Detailed Office Action, the Office acknowledges that Claim 36 is canceled and that Claims 32-35 and 37-56 are represented for examination.

Thus, the following argument (reproduced below from the present Office Action) is moot because claim 36 has been canceled:

"As to claim 36, Beshai teaches the invention as claimed, wherein the first and one or more subsequent transfer operations are performed in response to one or more RDMA (Remote Direct Memory Access) requests (see fig. 16 memory 0)"

Regarding rejections under 35 U.S.C. §103(a)

Claims 32-35 and 37-43 have been rejected under 35 U.S.C. §103(a) as being deemed unpatentable over Beshai et al. (U.S. Patent No. 6,721,271) in view of Massa et al. (U.S. Patent No. 6,658,469) and further in view of Edholm (U.S. Patent No. 6,600,721) and further in view of Colgate et al. (U.S. Patent No. 5,594,889).

Claims 44-47 have been rejected under 35 U.S.C. §103(a) as being deemed unpatentable over Beshai et al. (U.S. Patent No. 6,721,271), in view of Massa et al. (U.S. Patent No. 6,658,469), in view of Edholm (U.S. Patent No. 6,600,721) and further in view of Mauger (U.S. Patent No. 6,917,586) and further in view of Colgate et al. (U.S. Patent No. 5,594,889).

Claim 56 has been rejected under 35 U.S.C. §103(a) as being deemed unpatentable over Beshai et al. (U.S. Patent No. 6,721,271), in view of Massa et al. (U.S. Patent No. 6,658,469), in

view of Edholm (U.S. Patent No. 6,600,721) in view of Wang (U.S. Patent No. 6,477,612.) and further in view of Colgate et al. (U.S. Patent No. 5,594,889).

An embodiment of the Applicant's claimed invention improves efficiency of RDMA by coalescing (or merging) small RDMA requests into a single memory to memory transfer operation and dividing large RDMA requests into multiple operations to support hardware and software limitations. (*See*, for example, Page 15, lines 8-15 of the Applicants' application as originally filed.)

In an embodiment of the Applicants' claimed invention shown in Fig. 7, if an amount of data located in a first memory buffer (70) in a local system associated with a remote direct memory access (RDMA) request (62) does not exceed a maximum transfer size for a single memory to memory transfer operation over a data network to a remote memory buffer in a remote system, the data is associated with a first transfer operation (84) and if the amount of data associated with the first transfer operation (84) has not reached the maximum transfer size, a portion of the data (72A) associated with a subsequent transfer operation (86) for the RDMA request (62) located in one or more portions of one or more other memory buffers (72) in the local system is associated with the first transfer operation (84) to the remote memory buffer. (*See*, Fig. 7 and Page 15, line 16 – Page 20, line 9 of the Applicants' application as originally filed.)

To establish a prima facie case for obviousness under 35 U.S.C. 103(a), (1) there must be some suggestion or motivation to combine reference teachings; (2) there must be a reasonable expectation of success; (3) the references when combined must teach or suggest all the claim limitations. For the reasons discussed below, it is respectfully submitted that the Office has not established a prima facie case under 35 U.S.C. 103(a) for claims 32-35 and 37-55 and that therefore, claims 32-35 and 37-56 are allowable.

<u>The references when combined do not teach or suggest all the claim limitations</u>

The Office has cited Col. 4 lines 1-47 of Colgate reproduced below:

"FIG 4 is a flow chart of the memory resource allocation look ahead method when the evd.sub.-- get.sub.-- event routine is executed with the next.sub.-- size.sub.-- hint parameter. Shown in FIG. 4 are two co-operating algorithms, algorithm A and algorithm B, that are exchanging data. Algorithms A and B are shown in FIG. 4 as operating in separate portions of

memory and as having an address space boundary separating the two. As shown in FIG. 4, algorithm A starts out by allocating a memory block equal in size to a maximum predetermined size. Next, algorithm A calls algorithm B and asks algorithm B to fill this memory block. Further, as shown, algorithm B, upon receiving this request, first checks to see if it has more than 1 unit or pieces of data to transfer. If algorithm B has only 1 piece of data to transfer, it then appends a predetermined maximum size as the size of the next data piece to be transferred and transfers the first piece of data to algorithm A alongwith the appended size information in the first data transfer step. Also as shown in FIG. 4, if algorithm B has more than one piece of data to transfer it appends to the first data piece an estimate of the size of the second data piece to be transferred in the second data transfer step. Further as shown in FIG. 4, algorithm A uses the size information transferred with the first data piece to allocate a memory block corresponding to the appended size information of the subsequent data transfer from algorithm B to algorithm A. Further, as shown in FIG. 4, the size estimate appended with the first data piece could also be the actual size of the data to be transferred in the second data transfer step.
Further details of the event dispatch mechanism 16 are shown in FIG. 5. To accomplish its tasks, each of the event dispatcher's outbound streams 50 must form a long-term association with the corresponding target sink 18. The event dispatcher utilizes the services of a protocol engine 24 to do so. Each outbound stream may operate a different protocol, and therefore a different protocol engine 24.

It should be apparent to those skilled in the art that the system and method of this invention optimizes memory utilization when reading a stream of variable length records. By piggy-backing information concerning the memory requirements of a subsequent data transfer with a current data transfer operation, the invention allows memory to be allocated in closer correspondence to actual anticipated requirements, rather than in an arbitrary predetermined size. It should further be apparent to those skilled in the an that various changes in form and details of the invention as shown and described may be made. It is intended that such changes be included within the spirit and scope of the claims appended hereto."

as teaching:

"if the amount of data associated with the first transfer operation has not reached the maximum transfer size, associating a portion of the data associated with a subsequent transfer operation for the RDMA request located in one or more portions of one or more other memory buffers in the local system with the first transfer operation to the remote memory buffer."

as claimed by the Applicants in Claim 32.

Applicants cannot find any reference to the Applicants' claimed "associating a portion of the data ... with the first transfer operation" in the cited portion of Colgate. In contrast, Colgate merely discusses data transfer between algorithms operating in a memory, with the size of a subsequent data transfer being included with a first data transfer and the data transferred in the subsequent transfer. The size of a subsequent data transfer does not teach or suggest the Applicants' claimed "**a portion of the data**" that is associated "with the first transfer operation

to the remote memory buffer". In the system discussed by Colgate only the **size** of a subsequent transfer operation is transferred with the first data, the data that is associated with the subsequent data transfer is transferred in the subsequent data transfer.

Beshai has been cited for disclosing the Applicants' claimed "if an amount of data located in a first memory buffer in a local system...does not exceed a maximum transfer size for a single memory to memory...associating the data with a first transfer operation; and if the amount of data associated with the first transfer operation has not reached the maximum transfer capacity, associating data for the RDMA request located in one or more portions of other memory buffers with the first transfer operation". The Office admits that Beshai "does not explicitly disclose transfer operation over a data network to a remote memory in a remote system with other memory buffers in the local system associated with a remote data memory access (RDMA) request". (*See* Office Action mailed January 9, 2008, Page 3.)

Beshai discusses processing of received variable-sized network packets by dividing the packets into equal sized packet segments and aggregating packet segments destined for a particular egress module into a parcel. "[P]arcels are formed from an ingress buffer only when the number of segments in the buffer equals or exceeds the parcel size, q, or if the waiting time of the head-of-buffer segment has reached a predetermined threshold" (*See*, Beshai, col. 12, lines 34-38 (emphasis added).)

Beshai does not teach or suggest at least:

> "associating the data with a first transfer operation; and
> if the amount of data associated with the first transfer operation has not reached the maximum transfer size, associating a portion of the data associated with a subsequent transfer operation for the RDMA request located in one or more portions of one or more other memory buffers in the local system with the first transfer operation to the remote memory buffer."

as claimed by the Applicants in Claim 32.

In contrast to the Applicants' claimed invention, Beshai is merely directed to managing packet throughput within a 3-stage switch, by dividing packets into **predetermined equal size segments** and aggregating packet segments for a same egress module destination. "Segments are grouped into logical buffers according to egress-module destination". Beshai discusses

coalescing packet segments destined to an egress module in a switch in a parcel "only when the number of packet segments in buffer equals or exceeds the parcel size." (*See*, col. 12, lines 34 – lines 38.)

Beshai does not teach or suggest "associating a portion of the data associated with a subsequent transfer operation for the RDMA request located in one or more portions of one or more other memory buffers in the local system with the first transfer operation to the remote memory buffer". In contrast, Beshai merely discusses transferring packets received by ingress modules to egress modules within a switch based on "the number of parcels that each ingress module 32 is permitted to transfer to each egress module 36 during a specified transfer allocation period".

In contrast, data transferred within the switch from ingress modules to egress modules is based on packet segments of equal size. (*See,* Beshai, for example, Fig. 1, column 8, line 32 – column 9, line 9.) Beshai merely discusses aggregation of these equal size packet segments that are destined to egress from the same egress module into parcels with the number of packet segments transferred based on for example, "a nominal capacity of 100 packet segments per transfer allocation". (*See*, col. 15, lines 9-14.)

Massa has been cited for disclosing "transfer operation to a remote memory in a remote system with other memory buffers in the local system". However, Massa does not teach or suggest at least:

> "associating the data with a first transfer operation; and
> if the amount of data associated with the first transfer operation has not reached the maximum transfer size, associating a portion of the data associated with a subsequent transfer operation for the RDMA request located in one or more portions of one or more other memory buffers in the local system with the first transfer operation to the remote memory buffer."
>
> as claimed by the Applicants in Claim 32.

Massa's discussion of two methods of transferring data (RDMA for large data transfer and messages for small data transfer) teaches away from the Applicants' claimed method which coalesces small data transfers to transfer the maximum RDMA data in a single transfer operation. In contrast, Massa teaches a system that transfers multiple messages to transmit data that is not large enough to be transferred using an RDMA transfer operation. "The switch 126

detects the data size and decides whether to use RDMA data transfer or messages to transfer the data to application 132". (*See*, Massa col. 11, lines 63-67.)

Massa does not teach or suggest how much data is transferred in an RDMA transfer operation. Massa merely indicates that "remote switch 126 transfers an amount of data equal to the size of the set of receiving buffers 134 from the set of transmission buffers 138 into the set of receiving buffers 134 using one or more RDMA write operations". (*See*, Massa, col. 12, lines 48-52.) Thus, Massa merely describes a prior art RDMA system as discussed in the background of the Applicants' specification as originally filed, that is, "RDMA request can transfer data from multiple local memory buffers into a single remote memory region". (*See*, Applicants' specification, Page 2, lines 10-12.) In the RDMA system discussed by Massa, there is no teaching or suggestion of reducing the number of RDMA transfer operations to transfer data from local buffer(s) to a remote buffer by "associating a portion of the data associated with a subsequent transfer operation for the RDMA request located in one or more portions of one or more other memory buffers in the local system with the first transfer operation to the remote memory buffer" as claimed by the Applicants in claim 32 (emphasis added).

The Office has cited col. 4, lines 4-15 of Edholm for disclosing the Applicants' claimed "data is to be transferred in a single transfer operation to a remote memory buffer". The Applicants' cannot find any reference to a "single transfer operation" or "a remote memory buffer" in the cited portion of Edholm reproduced below:

"When a packet is to be transmitted from the protocol stack, the protocol stack assembles the packet, transfers the packet to a memory buffer and sets a pointer in the NDIS stack indicating the memory location stonng the packet. Driver 316 reads the pointer in the NDIS stack and when memory space is available on the NIC card 312, the data packet is read from memory to the NIC card. In order to implement pacing of the flow of data at the NDIS level, a control application 332 monitors when protocol stack 324 transfers a packet to the memory buffer or queue for transmission."

Edholm merely discusses a "memory buffer" and "NIC card" in a system such as a client system (*See,* for example, Fig. 2, 204). Cited reference Edholm merely discusses packet processing, that is, a method for transmitting packets over a network, between NIC cards 312 coupled to a network 308. There is no discussion of a "single transfer operation to a remote memory buffer". The Applicants' claimed invention relates to Remote Direct Memory Access (RDMA), that is, memory to memory transfer. One skilled in the art of RDMA would not look

to bandwidth management of a network connection between an end node and a switch to increase throughput in an RDMA application.

Claims 33-35 and 37-39 are dependent claims that depend directly or indirectly on claim 32 which has already been shown to be non-obvious over the cited art.

The Office has cited col. 10, lines 48-65 of Beshai as teaching the Applicants' claimed "associating a descriptor with the first transfer operation" as claimed in Claim 34. Col. 10, lines 48-65 of Beshai is reproduced below:

"FIG. 7 illustrates in more detail the core 64 of the rotator-based switch shown in FIG. 6. There are N core memories 66, each of which is logically partitioned into N sections 68, each section 68 being implicitly associated with an egress module and adapted to store a predetermined number K, 8 for example, of packet segments 44. Each inner link 38a that interconnects an ingress module 32 (not shown) to a rotator port (not shown) accesses each core memory 66 during a rotator cycle. The access duration is at least sufficient to transfer the predetermined number of packet segments K to the accessed memory section 66. The K packet segments 44 need not belong to the same packet 43 or to the same egress module 36, i.e., the packet segments 44 may be transferred to different sections 68 of the core memory 66. However, a given packet segment 44 can only be transferred to a memory section 68 that corresponds to an egress module 36 from which the packet 43, to which the packet segment belongs, is to egress from the switch 30."

Applicants cannot find any reference to the Applicants' claimed "descriptor", "remote memory buffer" or "RDMA request" in col. 10, lines 48-65 of Belshai as claimed by the Applicants in dependent claim 34. In contrast, Belshai merely describes "core memories 66" (for storing packet segments for a particular egress module in a switch. (*See*, col. 10, lines 48-65.) Belshai's discussion of "core memories" does not teach or suggest the Applicants' claimed "descriptor". (*See*, for example, Fig. 7, descriptor (80).) The Applicants' claimed descriptor specifies "the remote memory buffer to which the data is to be transferred and to indicate a portion of data remaining to be transferred for the RDMA request".

Independent claims 40, 44 and 48 recite a like distinction and are thus patentably distinguished over the cited art. Claims 41-43 depend directly or indirectly on claim 40, claims 45-47 depend directly or indirectly on claim 44 and claims 49-55 depend directly or indirectly on claim 48 and are thus patentably distinguished over the cited references.

Furthermore, the Office has cited col. 17, lines 23-27 of Wang as disclosing "the RDMA request is received via a single function call" as claimed by the Applicants in dependent Claim 56. Col. 17, lines 23-27 of Wang are reproduced below:

"24. The system of claim 23 wherein the memory is configured with a set of function calls to receive the allocation request for virtual memory, the allocation request or physical memory, and the request to map virtual memory to physical memory."

Applicants' are unable to find any reference to the Applicants' claimed "**RDMA** request" or "**single** function call" in col. 17, lines 23-27 of Wang. In contrast to the Applicants' "**single** function call", Wang merely discusses "a set of function calls".

There is no suggestion or motivation to combine reference teachings

Beshai has been cited for its teaching of packet management within a switch. Mauger has been cited for its teaching of "a host fabric adaptor". Edholm has been cited for its teaching of bandwidth management by adjusting latency between packets. Massa has been cited for its teaching of RDMA operations. Wang has been cited for its teaching of virtual memory management. Colgate has been cited for its teaching of associating a portion of data associated with a subsequent transfer operation. One skilled in the art of RDMA operations would not look to packet management within a switch, host fabric adaptors, adjusting latency between packets or virtual memory management to perform "a single memory to memory transfer operation over a data network to a remote memory buffer in a remote system" by "associating the data with a first transfer operation; and if the amount of data associated with the first transfer operation has not reached the maximum transfer size, associating a portion of the data associated with a subsequent transfer operation for the RDMA request located in one or more portions of one or more other memory buffers in the local system with the first transfer operation to the remote memory buffer".

Therefore, separately or in combination, Beshai, Edholm, Massa, Mauger, Wang and Colgate do not teach or suggest the Applicants' claimed invention. Even if combined, the present invention as now claimed does not result as argued above.

Thus, Applicants respectfully request that the rejection of Claims 32-35, 37-43 and 48-56 be withdrawn.

Accordingly, the present invention as now claimed is patentably distinguished from the cited references. Removal of the rejections under 35 U.S.C. § 103(a) and acceptance of claims 32-35 and 37-56 is respectfully requested.

P8193

CONCLUSION

In view of the foregoing, it is submitted that all claims (claims 32-35 and 37-56) are in condition of allowance. The Examiner is respectfully requested to contact the undersigned by telephone if such contact would further the examination of the above-referenced application.

Should an extension of time be necessary to respond to the outstanding Office Action, applicants respectfully petition for an extension of time pursuant to 37 C.F.R. § 1.136(a). Please charge our Deposit Account No. 50-4238 to cover the fee for the extension.

Respectfully submitted,

Customer Number: 59796

Date: 5/9/2008

/Caroline M. Fleming/
Caroline M. Fleming
Reg. No. 45,566
Telephone No. (978) 553-7371

CERTIFICATE UNDER 37 CFR 1.8: The undersigned hereby certifies that this correspondence is being deposited via facsimile addressed to: MS Amendment, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450, on this 9 day of May, 2008.

Kyrstin Ryan
Name

/Kyrstin Ryan/
Signature